Published in final edited form as:
J Exp Psychol Gen. 2014 December ; 143(6): 2316–2329. doi:10.1037/xge0000022.

Stimulus devaluation induced by stopping action

Jan R. Wessel[1,*], **John P. O'Doherty**[2], **Michael M. Berkebile**[1], **David Linderman**[1], and **Adam R. Aron**[1]

[1]Department of Psychology, University of California, San Diego

[2]Division of Humanities and Social Sciences and Computation and Neural Systems Program, California Institute of Technology, Pasadena, CA

Abstract

Impulsive behavior in humans partly relates to inappropriate overvaluation of reward-associated stimuli. Hence, it is desirable to develop methods of behavioral modification that can reduce stimulus value. Here, we tested whether one kind of behavioral modification – the rapid stopping of actions in the face of reward-associated stimuli – could lead to subsequent devaluation of those stimuli. We developed a novel paradigm with three consecutive phases: implicit reward learning, a stop-signal task, and an auction procedure. In the learning phase, we associated abstract shapes with different levels of reward. In the stop-signal phase, we paired half those shapes with occasional stop-signals, requiring the rapid stopping of an initiated motor response, while the other half of shapes was not paired with stop signals. In the auction phase, we assessed the subjective value of each shape via willingness-to-pay. In two experiments, we found that participants bid less for shapes that were paired with stop-signals compared to shapes that were not. This suggests that the requirement to try to rapidly stop a response decrements stimulus value. Two follow-on control experiments suggested that the result was specifically due to stopping action rather than aversiveness, effort, conflict, or salience associated with stop signals. This study makes a theoretical link between research on inhibitory control and value. It also provides a novel behavioral paradigm with carefully operationalized learning, treatment, and valuation phases. This framework lends itself to both behavioral modification procedures in clinical disorders, and research on the neural underpinnings of stimulus devaluation.

Keywords

Inhibitory Control; Value; Devaluation; Cognitive Control; Stop Signal Task

INTRODUCTION

Overvaluation of reward-associated stimuli (e.g., a cigarette or unhealthy food) is a core symptom of psychiatric disorders such as substance dependence (Goldstein & Volkow, 2002, 2011), unhealthy eating (Rothemund et al., 2007), and pathological gambling (Clark et al., 2013; Shao, Read, Behrens, & Rogers, 2013). One way to reduce such valuations is

CORRESPONDENCE: Jan R. Wessel, PhD, 9500 Gilman Drive, #0109, University of California, San Diego, La Jolla, CA 92093, USA, Phone: 858 822 2129, Fax: 858 534 7190, jwessel@ucsd.edu.

through behavior modification. Several methods have been tried, such as cue exposure therapy (Hammersley, 1992; Heather & Bradley, 1990), the retraining of approach tendencies (Fishbach & Shah, 2006; Wiers, Eberl, Rinck, Becker, & Lindenmeyer, 2011) and choice-induced preference change (Izuma et al., 2010). However, such interventions have only achieved mixed success (Conklin & Tiffany, 2002; Dutra et al., 2008; Marteau, Hollands, & Fletcher, 2012). A different way to change value could perhaps be achieved by requiring people to stop a motor response in the face of the valuable stimulus. This could be effective for several reasons. Motor stopping might be a form of 'avoidance', which leads to changes of explicit or implicit attitudes (Phaf & Rotteveel, 2012). Alternatively, repeatedly stopping towards a stimulus could 'tag' the stimulus with an inhibitory signal (Lenartowicz, Verbruggen, Logan, & Poldrack, 2011; Verbruggen & Logan, 2008). A third alternative is that rapid action-stopping could have suppressive effects not just on the motor level, but also on cognition (Wessel, Huber, & Aron, 2012) – which could include value representations. Finally, it could also be the case that motor stopping reduces value through motor/limbic 'spill-over' (Berkman, Burklund, & Lieberman, 2009).

Recent studies have taken several approaches to investigate whether motoric stopping induces stimulus devaluation. These have mostly used Go/NoGo or stop-signal paradigms along with primary reinforcers such as palatable foods (Houben, 2011; Houben & Jansen, 2011; Veling, Aarts, & Papies, 2011; Veling, Aarts, & Stroebe, 2013a, 2013b) or alcoholic beverages (Houben, Havermans, Nederkoorn, & Jansen, 2012; Houben, Wiers, & Jansen, 2011). Such studies have asked, for example, if pairing a stimulus (such as a picture of chocolates) with action-stopping (NoGo-trials or stop trials) leads to subsequent reductions in consumption. Other studies have reported that action-stopping leads to the emotional 'devaluation' of faces (as measured by their trustworthiness; Doallo et al., 2012; Fenske, Raymond, Kessler, Westoby, & Tipper, 2005; Kiss et al., 2007), the reduction of motivational value of sexual stimuli (Ferrey, Frischen, & Fenske, 2012), and changes in brain activity for emotional stimuli (Berkman et al., 2009). A related study showed that when participants were in a state of motor caution (i.e., they expected stop signals to occur) they took fewer risks in a gambling task (Verbruggen, Adams, & Chambers, 2012).

However, it is still not clear what causes these effects. The reductions in value or motivation could reflect so-called 'task demand characteristics', wherein participants consciously or subconsciously perform according to an experimenters' expectations (especially when the purpose of a study is obvious). The reductions in value could also reflect, as noted above, a variety of possible mechanisms, which include inhibitory tags and attitude change. Here, we set out to clearly demonstrate that motor stopping itself induces an actual change in stimulus valuation.

Key challenges for any study of stopping-induced stimulus-devaluation are proper operationalization of the stopping intervention and proper assessment of subjective value. A 'true' action-stopping situation requires subjects to first initiate a go-response and then to stop that response when a signal occurs, as is achieved with the stop signal task (SST; Logan, Cowan, & Davis, 1984). If the go-response is highly prepotent (because it must be made very fast and/or because the proportion of go trials is high and the proportion of stop-trials is low) participants will have to stop an initiated response every time (instead of

simply making a decision not to respond in the first place), and this forces them to engage an active mechanism of inhibitory control (Aron, 2007). Yet, many of the above-mentioned studies of devaluation have instead used the Go-NoGo task, furthermore often with equal proportions of Go and NoGo trials. Here, one stimulus (e.g., a face) is always paired with the go-response requirement, whereas another is paired with the NoGo-requirement. This task set-up is less like a response inhibition paradigm and more like a decision-making paradigm. Because the Go response is minimally prepotent and there is no speed pressure, participants can perform the task by deciding whether to Go or not, rather than Going on every trial and then having to rapidly stop the Go response. Here, we aimed for a task design in which outright action stopping was definitely engaged.

Another important element for a paradigm that aims to test whether motor stopping induces stimulus devaluation is a careful behavioral operationalization of value. Many of the above-mentioned studies of devaluation use an inherently valuable good (such as chocolate or water), perform the action-stopping treatment, and then measure value indirectly, either through the amount of consumption of the good (Houben, 2011), self-report likert scales (Veling et al., 2013a), or through potentially related constructs such as judgments of ‘trustworthiness’ (Doallo et al., 2012; Fenske et al., 2005; Kiss et al., 2007). However, it is problematic to use objects with pre-existing valuations and also problematic to measure value through consumption or self-report. First, such experiments can be quite transparent in design, making them prone to the above-mentioned task demand characteristics. For example, if a Go stimulus is paired with one kind of drink stimulus, and a NoGo stimulus with another, the participant could soon realize the experimenter’s intention and behave accordingly. Second, using pre-established value makes it difficult to have value levels under proper experimental control, as the value of an object or good will differ drastically between participants, and even within participants across time.

The Current Study

In order to test whether motor stopping induces stimulus devaluation, we designed a novel behavioral paradigm that included both a proper operationalization of stopping, as well as a standard behavioral economics assessment of subjective value. Our experimental framework had three phases. In the first phase (learning), neutral artificial stimuli (geometric shapes) were implicitly associated with different levels of monetary reward, ensuring tight control over the initial level of stimulus valuation. The second phase was the ‘treatment’. In Experiment 1, subjects performed a stop-signal task (SST) with a dynamically adjusted stop-signal delay, ensuring a race-like stopping-process, which requires ‘true’ action-stopping on stop-trials. Crucially, half of the shapes from the learning-phase were paired with occasional stop-signals in this treatment phase, while the other half of the shapes were never paired with stop signals. In the third phase (value measurement), we assessed subjective valuations of the shapes using an auction-procedure that measures willingness-to-pay. This is a standard way to assess value that is routinely used in behavioral economics (Becker, Degroot, & Marschak, 1964). For Experiment 1, we predicted that participants would bid lesser amounts of money for shapes that were paired with stop-signals in the treatment phase compared to shapes that were not. Experiment 2 aimed to replicate those results. Experiments 3 and 4 were control experiments in which the treatment phase did not include

action-stopping, but instead aimed to investigate the influence of other variables on stimulus valuation, such as conflict- or error processing, stimulus aversiveness, effort, and attentional capture. All these factors could potentially play a role in stimulus devaluation observed in a stop-signal task, but are not directly related to actual action-stopping. Hence, we aimed to test whether those factors themselves could account for stimulus devaluation; or whether the decisive factor is action stopping itself.

EXPERIMENT 1

Method

Participants—36 participants at UCSD provided written informed consent and received a $10 base-payment, plus money earned in the task. Three participants were excluded for not understanding the auction-phase. One participant was also excluded when we discovered he had been informed about the objective of the experiment prior to participation, leaving N = 32 (mean age = 21.5y, SEM = .63, range = 17 – 31; 22 female; 2 left-handed).

Materials—Stimuli were presented on Apple Macintosh computers (Apple Inc., Cupertino, CA) running Psychtoolbox 3 (Brainard, 1997) on Matlab 2009b (TheMathWorks, Natick, MA). Responses were made using a QWERTY-keyboard.

Experimental Task—The task was divided into three parts (Figure 1).

Learning-phase: In this phase, we associated 8 different geometric shapes with four different monetary values. Two shapes each were associated with mean values of $.5, $1, $2, or $4, respectively. The shapes were a square, a circle, a diamond, a triangle, an inverted triangle, a cross, a hexagon, and an “I”-shape (see Figure 1). The different shapes were also colored (white, green, blue, yellow, cyan, magenta, orange, or gray), in order to maximize the chances of acquiring an implicit value association. Color-shape pairings remained constant throughout the experiment. The shapes were randomly assigned to a given value for each participant.

A trial proceeded as follows. A fixation cross was presented for 500ms, which divided the screen into four quadrants. Then, one of the eight shapes appeared in one of the quadrants. Participants pressed one button (out of four possible) corresponding to that quadrant. There was a deadline of 1000ms; if no response was made within that period, a “too slow” message appeared for 1000ms. Immediately after the response, a payment amount was indicated on the screen: a black square was superimposed on the stimulus display, within which the message “You won $ X.XX” appeared (for 1000ms). Participants were told that responding “as fast and accurately as possible” would lead to higher rewards. To ensure that participants paid attention to the magnitude of reward in this phase, we told them that the computer would randomly pick five of the trials at the end of the experiment and pay out the amount associated with those trials. After the reward display, the screen was cleared before the next trial, resulting in a trial duration of 3000ms. Trials on which misses (no response before deadline) or errors (wrong button) occurred were repeated until the correct response was made. Most importantly, unbeknownst to the participants, the trial-by-trial monetary reward schedule was actually independent of performance, and instead followed a pre-

defined schedule: Each stimulus had one of four fixed distributions of possible payout values randomly assigned to them. The distributions centered around four different means (either $.5, $1, $2, or $4), with a uniform dispersion of +/− 25c around those means. To ensure that the value representations were largely implicit, only 80% of the trials per shape were rewarded according to this distribution. On the remaining 20% of trials, the reward was 0[1]. Each shape had an identical probability of appearing in any of the four quadrants. Participants performed 50 trials per shape (400 overall), which were presented in 4 blocks.

Treatment-phase: In this phase, we used the 8 shapes from the earlier learning phase, which were now associated with different values. The primary task for the participants was to make a quick motor response according to the placement of the shape on the computer screen (left or right), and to stop their impending response when a stop-signal occurred. Four out of these eight shapes (one per value-step, i.e., one shape each associated with a mean value of $.5, $1, $2, or $4) were always paired with Go trials (i.e., a stop signal was never presented on trials with these shapes). The other half of the eight shapes (which were also associated with mean values of $.5, $1, $2, or $4), were paired with stop signals on some of the trials (on 76.4% of trials for each stimulus; i.e., 26 out of 34 trials). Ideally, these latter shapes would be paired 100% of the time with stop signals (to increase the potential effect of action stopping on stimulus valuation). However, a relative probability of stop vs. go-trials of .5 would diminish the prepotency of the go-response. Furthermore, we were concerned that a fully deterministic pairing of some shapes with stopping on 100% of trials would potentially lead to the emergence of explicit knowledge on the participants' part. Such awareness of the pairing might cause participants to consciously withhold their response on these shapes on trials after they picked up on the contingency, which would turn this task into a decision making task instead of a stopping-task (because participants might just decide to never initiate a response once they realize that some shapes are always paired with stopping). Hence, we decided to pair the stopping-shapes with stopping on 76.4% of the trials only (resulting in an overall probability of a stop-signal of around 38%), in order to achieve a good tradeoff between maximizing the potential effects of stopping on devaluation and avoiding the detrimental effect of a deterministic contingency.

A trial proceeded as follows. The screen was divided into two halves by a vertical line for 500ms. The same shapes from the learning phase then appeared to the left or right of this line. Participants were instructed to respond as fast and accurately as possible according to the position of the shape by using one of two buttons on the keyboard (deadline: 1000ms), one with their left hand, and one with their right hand. Participants were instructed that occasional stop-signals (200ms sine-wave tones, 900Hz) would occur shortly after stimulus-onset – in which case they should try to cancel the response. The stop-signal delay (SSD) was adapted separately for left and right responses depending on ongoing performance (+50ms following successful stop-trials, −50ms following failed stop-trials) to achieve an overall probability of successful stopping p(stop) of .5 (Verbruggen & Logan, 2009). The

[1]The fact that shapes were only rewarded according to the predetermined schedule on 80% of the trials actually reduces their associated value to .8 * [$.5, 1, 2, 4]; i.e., the actual values associated with the stimuli are $.4, $.8, $1.6, and $3.2, respectively. For simplicity reasons, we will continue to refer to the value steps according to the means of their reward distributions, i.e., $.5, $1, $2, and $4.

SSD's initial value was set to 250ms. Participants were instructed that successful stopping on stop-trials and fast responding on go-trials were equally important. Participants performed 34 trials for each shape (272 overall), split into four blocks. In the breaks between blocks, participants received information about their Go-trial reaction time (GoRT), as well as their miss- and (direction-) error-rates. Additionally, the experimenter (but not the participant) received information about p(stop) and SSD to ensure that participants were not overly favoring stopping over going or vice-versa. In case the participant appeared to favor either strategy, the experimenter informed the participant to remember that both stopping and going were equally important. We aimed to achieve the following behavioral parameters in each block: GoRT between 400 and 650ms, p(stop) between .4 and .6, and SSD > 100ms. Note that in this treatment phase subjects were not reimbursed – in effect, the cues were presented in extinction.

Valuation-phase: In this phase, we repeatedly presented each of the 8 shapes (the same ones as those from the earlier two phases) to the participants and instructed them to bid money on each shape according to an auction procedure (Becker et al., 1964) designed to assess the 'true' subjective value of a given good or object. The participants were presented with six different cent-amounts that they could chose as their bid using one of six keys on the keyboard (r, t, y, u, i, o; the button-mappings were spatially congruent and displayed on the screen under the respective values, see Figure 1). Participants were instructed to choose the cent-amount that most closely represented their subjective valuation of the shape. This was implemented using an auction procedure (the exact instructions can be found in the Appendix, as well as on www.aronlab.org/Pubs/BDM.pdf; see Becker et al., 1964 for the general method). Participants were instructed to neither overbid nor underbid, as both of those strategies would be suboptimal. They were told that optimal bidding behavior would lead to larger payoffs, and that the payoff from the auction phase would be paid out at the end of the experiment, in addition to the payment from the initial learning phase.

A trial proceeded as follows. A fixation-cross was displayed for 500ms, after which one of the shapes from the earlier two phases was presented centrally for 1,500ms. Then, the six potential bids and their button mappings were presented, and participants had 5,000ms to pick an amount to bid. The overall trial-duration was fixed at 6,000ms, with the remainder of the 5,000ms response-window going to the ITI. Each of the 8 shapes was presented ten times (i.e., there were 80 trials total), each time with six potentials bids, as explained above. These bids came from five different sets of values. Specifically: [34, 68, 102, 136, 170, 204], [39, 78, 117, 156, 195, 234], [44, 88, 132, 176, 220, 264], [49, 98, 147, 196, 245, 294], and [54, 108, 162, 216, 270, 324]. We chose to use multiple different sets of values in order to induce variance into the bidding behavior, specifically, so that participants did not bid the exact same value every time they saw a given shape. These sets of bids were chosen so that the range of values completely covered the range of true values that were associated with each shape in the learning phase (40, 80, 160, and 320 cents, respectively). Each set was presented twice for each shape, in random order. The bids within each set were randomly assigned to one of the six response buttons on each trial. The shapes themselves were also presented in random order across the 80 trials. Trials were presented in 4 blocks.

Procedure—Participants were first instructed on both the learning-phase and the treatment-phase. They practiced both phases briefly (10 trials each), before performing the actual learning- and treatment-phases. Thereafter, they read the instructions of the auction. The experimenter made sure that the auction procedure was fully understood, placing special emphasis on the fact that systematic over- or underbidding, or bidding according to any other rationale than the participant's subjective value, would lead to suboptimal earnings at the end of the experiment. After the valuation-phase, participants were debriefed. We used a questionnaire to verbally probe the amount of explicit knowledge about the two main regularities of the procedure: 1) the reward schedule in the learning phase [question 1 below] and 2) the fact that only some shapes were paired with stop-signals in the stopping phase [question 2 below]. The experimenter asked the participants the following five questions [numbering is for the reader and was not provided to participants]:

1a) What strategy did you use in the last part of the experiment? [auction procedure]

1b) Did you think that any shapes were associated with higher or lower reward during the first phase of the experiment? [learning phase]

1c) If so, which? If so, did you utilize that information in the last phase? [auction phase]

2a) Did you notice anything about the second phase of the experiment? [stopping phase]

2b) Did you notice that some shapes were paired with stop signals more often than others? If so, which? [stopping]

3 Any other comments or remarks?

Finally, the computer picked 5 trials from the learning-phase, which were added and paid out to the participants, in addition to a $2 bonus for the valuation-phase (while participants were bidding with the expectation that five random trials would be paid out, we did not actually perform an artificial auction, and instead paid this constant amount for simplicity). Participants came away from the experiment with between $4 and $14 in addition to the base payment of $10.

Analysis

Valuation-phase: The main effects of interest were in the valuation-phase. These were tested using repeated-measures analyses-of-variance (rmANOVA) with the independent variables VALUE (4 levels) and STOPPING (2 levels), and the dependent variable BIDDING-LEVEL (1–6, from the lowest option within a set of bids to the highest).

Treatment-phase: The SST data were also analyzed to ensure the validity of the race model and to verify that stopping was done in a typical way. To examine the validity of the race model (Verbruggen & Logan, 2009), we tested whether RT on go-trials was slower than RT on failed stop-trials in each participant's data. We also examined stop signal performance both in terms of the probability of stopping (which should be in the range .4 to .6) and stop-signal reaction time (SSRT), which should be in the range of about 130 to 300ms, based on

typical manual stopping paradigms using auditory stop-signals. We calculated SSRT using the mean method (Verbruggen & Logan, 2009).

Post-hoc assessment of explicit knowledge: The questionnaire assayed the level of explicit knowledge of the two main regularities of the procedure: 1) the reward schedule in the learning phase and 2) the fact that only some shapes were paired with stop-signals in the stopping phase. The classification was done as follows:

1. Explicit learner. This designation was given if participants spontaneously verbalized one of the regularities in response to the 'open' questions of whether they noticed something about the first or second phase of the experiment (1a and 2a), and displayed significant accurate knowledge about the regularity. They needed to accurately name the relative values of more than half of the stimuli to be labeled as an explicit learner of the learning regularity. They needed to accurately name all four shapes that were either paired with stopping or never paired with stopping to be labeled as an explicit learner of the stopping regularity.
2. Partially Explicit Learner. This designation was given if a participant reported a 'feeling of knowing' of one of the regularities in response to the 'open' questions about the learning and stopping phases (questions 1a and 2a, respectively), and could accurately name at least one shape for which the regularity was true. For example, for the learning phase (question 1), he/she could state "The green square was always paired with high reward", or "The white diamond was always paired with low rewards". With regards to the stopping phase (question 2), he/she could state e.g., "The blue circle was often paired with stopping", or "The yellow triangle was never paired with stopping".
3. Implicit learner. These participants reported not noticing any regularities in response to any of the questions in the debriefing questionnaire.

To avoid participants' realizing the purpose of the experiment in the treatment phase (i.e., task demand characteristics), we excluded all participants that had partial explicit knowledge of the stopping contingency.

Results

Treatment phase—Based on the debriefing (see below), four participants were classified as explicit or partially explicit learners of the stopping-contingency and were excluded from the analysis (in all four of these participants, the average SSD was very long, and SSRT estimates were < 100ms, which is unrealistic, and speaks to a 'waiting strategy', i.e., a go-response is never initiated). Mean GoRT in the remaining sample was 594ms (SEM: 20ms), failed-stop RT was 499ms (SEM: 19ms), and this difference was significant ($t(27) = 14.02$, $p < .0001$, $d = .95$), validating the independence assumption of the race model. Error- (.05%) and miss-rates (2.76%) were low. Mean p(stop) was .56 (SEM: .01); SSD was 423ms (SEM: 20ms), SSRT was 171ms (SEM: 7ms).

Valuation phase—The 2×4 rmANOVA revealed a significant main effect of VALUE ($F(3/81) = 6.11$, $p < .001$, partial $\eta^2 = .18$), showing that participants bid more for shapes

that were associated with higher rewards in the learning phase. There was also trend towards a main effect of STOPPING ($F(1/27) = 3.05$, $p = .09$, partial $\eta^2 = .1$), showing that participants tended to bid less on shapes that were paired with stop-signals in the treatment phase. There was no interaction between VALUE and STOPPING ($F(3/81) = .83$, $p = .48$). Descriptively, 17 out of the 28 participants showed stopping-related devaluation (measured by a difference of the mean bidding level for all non-stopping-shape bids and the mean bidding level for all stopping-shape bids). However, this measurement was contaminated by four outliers (criterion: × > 1.5*interquartile range), two in each direction, which invalidate the assumption of a normal distribution underlying parametric rmANOVA. When excluding these outliers from the rmANOVA, the main effect of VALUE was still significant ($F(3/69) = 6.41$, $p < .0001$, partial $\eta^2 = .22$), but now the STOPPING main effect was also significant ($F(1/23) = 7.48$, $p = .012$, partial $\eta^2 = .25$, Figure 2A), showing that participants bid less for shapes that were paired with stop-signals compared to those that were never paired with stopping. There was no interaction ($F(3/69) = 1.22$, $p = .3$). Additionally, non-parametric testing of the devaluation difference scores (mean bidding level for stopping shapes vs. mean bidding level for non-stopping shapes) including the outliers also revealed significantly reduced bids for the stopping-shapes (Wilcoxon's signed-rank test, $z = 1.7$; $p < .05$; one-sided).

Debriefing questionnaire—Out of the initial 32 participants, 4 were categorized as partially explicit learners of the stopping-contingency and were excluded, as stated above. Of the remaining 28 participants, none were categorized as explicit learners of the reward contingency. Nine were partially explicit learners of the reward contingency. The amount of explicit knowledge was limited, however. Out of the 9 participants, 7 could not accurately name more than one shape that was noticeable high or low in value, and none could name more than 3.

Discussion

We designed a novel experiment with implicit learning, treatment, and valuation phases. The treatment phase involved a stop-signal procedure that required motor stopping (rather than Go/NoGo decision making), and the valuation phase involved an auction procedure that assays value in an economically tractable fashion. In accordance with our prediction, we found that participants bid less for stimuli in the auction phase that had been paired with stop signals in the treatment phase than if they had not. This result suggests either that motor stopping itself or something about the stop-signal procedure is effective in reducing stimulus value.

Unfortunately, in this experiment, some participants used a 'waiting strategy', probably owing to their explicit knowledge of the stopping contingency, and hence were excluded from further analyses. Although the remaining participants did not report any partial explicit knowledge of the stopping contingency and had normal range values for the probability of stopping and SSRT (suggesting they did indeed stop initiated responses and did not 'wait out' their responses), we nevertheless aimed in Experiment 2 to replicate our result that stopping reduces value, now using a stop-signal treatment that was less susceptible to these effects. Now, the treatment-phase was split into eight blocks instead of four, which gives the

experimenter more opportunity to potentially reinstruct participants to not overly favor stopping over going, and which also diminishes the likelihood of explicit knowledge, as fewer trials are performed in immediate succession.

Experiment 2

Method

Participants—27 participants (mean age = 19.3y, SEM = .29, range = 18 – 23; 14 female; 3 left-handed) at UCSD provided written informed consent and received a $10 base-payment, plus money earned in the task. These were different participants from Experiment 1.

Materials, task, procedure, statistics—These were identical to Experiment 1, except that the treatment-phase was divided into eight blocks instead of four, and the number of presentations per shape was increased from 34 to 36, in order to bring the stop-signal probability on shapes paired with stopping to exactly 75% (27 of 36 trials).

Results

Treatment phase—Mean GoRT was 498ms (SEM: 15ms), failed-stop RT was 421ms (SEM: 14ms), and this difference was significant ($t(26) = 14.48$, $p < .0001$, $d = 1.03$), validating the independence assumption of the race-model. Error- (.2%) and miss-rates (.93%) were low. Inspection of the stopping parameters (SSD, SSRT, p(stop)) and verbal reports indicated that none of the participants now used a waiting strategy; hence all remained in the sample. p(stop) was .5 (SEM: .01), SSD was 332ms (SEM: 19ms), SSRT was 166ms (SEM: 6ms).

Valuation phase—The 2×4 rmANOVA revealed a marginally significant main effect of VALUE ($F(3/78) = 2.65$, $p = .052$, partial $\eta^2 = .09$) and no main effect of STOPPING ($F(1/26) = 2.22$, $p = .15$, partial $\eta^2 = .08$), with no interaction ($F(3/78) = .09$, $p = .97$). However, descriptively, 20 out of the 27 participants showed stopping-related devaluation. Again, this measurement was contaminated by outliers (two in each direction, four overall). After excluding these outliers from the rmANOVA the VALUE main effect was still significant ($F(3/66) = 2.11$, $p = .1$, partial $\eta^2 = .09$), and now the STOPPING main effect was also significant ($F(1/23) = 6.96$, $p = .015$, partial $\eta^2 = .24$, Figure 2B), replicating the finding from Experiment 1; the interaction was not ($F(3/66) = .3$, $p = .83$). Additionally, as for Experiment 1, non-parametric testing of the devaluation difference scores including the outliers also revealed significantly reduced bids for the stopping-shapes (Wilcoxon's signed-rank test, $z = 2.1$; $p < .05$).

Debriefing questionnaire—One of the participants reported explicit knowledge of the stopping-contingency. However, this participant could not name a shape that was regularly paired with stopping, or alternatively, a shape that was never paired with stopping. Furthermore, SSRT was within normal range. Hence, those data remained in the sample.

None out of the 27 participants explicitly learned the reward contingency in the learning phase. 6 out of the 27 participants were classified as partial explicit learners. Out of the 6

participants, 4 could not correctly name more than one shape that was noticeably high or low in value, and none could name more than 4.

Role of explicit knowledge—Post-hoc, it was of interest whether the presence or absence of (partial) explicit knowledge of the reward contingency had any influence on the magnitude of the putative stopping-induced devaluation. For example, one might expect that an explicit value representation is too 'strong' to be susceptible to devaluation (Marteau et al., 2012). Hence, we pooled data across Experiments 1 and 2 for the 15 participants who were classified as partial explicit learners of the reward regularity. We reran the ANOVA as above. Interestingly, there was a main effect of STOPPING in this sample ($F(1/14) = 16.36$, $p = .0012$, partial $\eta^2 = .54$, Figure 3). This result runs contrary to the intuition that explicit value representations would be less susceptible to devaluation; here the participants with some explicit knowledge also showed a stopping-induced devaluation effect; and moreover the effect size was even bigger compared to the full samples. We return to the implications of this in the General Discussion. While these participants thus had some explicit knowledge of which shapes were high vs. low value, none of them reported explicit or partially explicit knowledge of the stopping regularity. This is important because it strongly suggests the core result, that these participants bid less for stimuli that were paired with stop signals than those that were not, was *not* due to their adjusting performance according to what they thought the experimenter wanted regarding the stopping treatment (i.e., task demand characteristics). Instead, the results suggest that it was the stopping treatment that reduced value.

Discussion

We reran our novel paradigm with a minor change to the treatment phase. This phase now had 8 blocks instead of 4, giving the experimenter more opportunities to correct behavior in case participants employed a waiting-strategy on stopping-shapes, and helping to prevent participants from developing explicit knowledge of the stopping contingency. Experiment 2 now produced ideal stopping behavior. Crucially, it confirmed the key finding of Experiment 1: participants bid less for stimuli that had been paired with stop signals than those that had not. The effect size was comparable for both experiments (partial $\eta^2 = .25$ and .24), which is large as per common convention (Stevens, 2009). This greatly increases our confidence in the basic effect.

In the following Experiments, we aimed to elucidate the exact underlying mechanism by which stimulus devaluation is induced by the stop-signal treatment. Specifically, apart from action stopping itself, stop-trials differ from go-trials in the following ways, which are not necessarily related to stopping, but could potentially account for stimulus devaluation:

1. Aversiveness of the tone: on stop-trials, there is a tone (the stop-signal), which could be aversive.
2. Attentional capture: because the tone is relatively rare in the context of the overall treatment phase (it happens on 37.5% of trials overall, and only occurs on stop-shapes), it could capture attention.

3. Error rate: because the SSD is titrated in order to produce a p(stop) of .5, participants fail to stop about 50% of the time, i.e., they make commission errors on about 37.5% of all trials on the stopping-shapes.
4. Effort: stopping is potentially more effortful than going.
5. Conflict: stopping and going are incompatible responses, which could potentially lead to response-conflict (Botvinick, Nystrom, Fissell, Carter, & Cohen, 1999).

Hence, we designed two control experiments to examine the influence of the above factors on value. The general framework was identical to Experiments 1 and 2: An initial learning phase was followed by a treatment phase and a final auction phase. The learning and auction phases were themselves identical to Experiments 1 and 2; we only changed the treatment phase.

In the first control experiment (Experiment 3), we used a Simon paradigm (Simon & Rudell, 1967), in which an incongruent spatial mapping between an imperative stimulus and its mapped motor response leads to response-conflict, greater error rates, and higher effort in responding. By selectively pairing such incongruent trials with one half of stimuli only (75% of the time), we aimed to achieve a selective pairing of one half of stimuli with increased error commission, increased response-conflict, and increased effort, but importantly, not with motor stopping.

In a second control experiment (Experiment 4), we used a double-response paradigm (Verbruggen, Aron, Stevens, & Chambers, 2010), which uses the same tones that we used in Experiments 1 and 2 for stop-signals, but with a different instruction. In this paradigm, whenever a tone occurred (which happened at stimulus onset, i.e., with an interval of 0ms), participants had to first execute the primary response (press the button according to the side on which the stimulus was displayed), and then press a second button to indicate that they heard the tone. Here, double-response trials are different from single-response trials in that there is a tone (the same tone that was the stop-signal in Experiments 1 and 2), which could be aversive or capture attention. Furthermore, responding is more effortful than in the single-response condition, in that two responses have to be made instead of one. Again, identical to the logic of Experiments 1–3, we paired double-response trials with one half of the stimuli only, on 75% of trials for those stimuli.

Between these two control experiments, all the above stopping-unrelated factors that differentiate stop-trials from go-trials in the SST are engaged (aversiveness, response-conflict, salience, effort, error commission). If stimulus-devaluation in our Experiments 1 and 2 was indeed due to action-stopping, then neither Experiment 3 nor Experiment 4 should reveal stimulus-devaluation for shapes paired with incongruent Simon trials (Exp. 3) or with double-responses (Exp. 4).

Experiment 3

Method

Participants—26 participants (mean age = 20.46y, SEM = .33, range = 18–25; 24 female; 3 left-handed) at UCSD provided written informed consent and received a $10 base-

payment, plus money earned in the task. These were different participants from Experiments 1 and 2.

Materials, task, procedure, statistics—These were identical to Experiments 1 and 2, except for the treatment-phase. In the treatment phase, there were no stop-signals; instead participants performed the Simon task. Rather than being instructed to respond according to the side of the screen (Experiments 1 and 2), they were now instructed to respond according to an outer polygon enclosing the shape. The outer polygon could either be a gray rectangle or an equally sized gray diamond (Figure 4). For each subject, one of these polygons was mapped to the left response button, and the other to the right response button (response buttons were identical to Exp. 1 and 2; order counterbalanced across subjects). Response mappings were continuously displayed on the bottom of the screen. For one half of the shapes, the side of the response and the side of the stimulus (shape and outer polygon) were always identical (congruent trials). For the other half of shapes, the side of the response button was opposite the side of stimulus presentation on 75% of trials (27 out of 36). Participants were instructed to respond as quickly as possible according to the outer polygon. Trial timings, trial numbers, and block numbers were identical to Experiments 1 and 2.

Results

Treatment phase—The treatment phase was successful in evoking the classic pattern of results for Simon paradigms. Mean RT for congruent stimuli was 446ms (SEM: 8.9); mean RT for incongruent stimuli was 498ms (SEM: 10.1). Every subjects' incongruent RT was slower than their congruent RT (range: 20 – 111ms), and this difference was significant on the group level ($t(25) = 12.04$, $p < .0001$, $d = 1.1$). Error rate for congruent stimuli was 1.88% (SEM: .36); error rate for incongruent stimuli was 8.3% (SEM: .63). Every subjects' incongruent error rate was higher than their congruent error rate (range: 2% – 13.3%), and this difference was significant on the group level ($t(25) = 10.68$, $p < .0001$, $d = 2.49$). Miss rates were low across both trial types (.47% for congruent and .57% for incongruent trials).

Valuation phase—The 2×4 rmANOVA revealed a significant main effect of VALUE ($F(3/75) = 3.7$, $p = .015$, partial $\eta^2 = .13$) and no main effect of TREATMENT ($F(1/25) = 1.25$, $p = .27$, partial $\eta^2 = .048$), with no interaction ($F(3/75) = .07$, $p = .97$). One outlier contaminated the devaluation measurement. Excluding this outlier from the rmANOVA did not change the VALUE main effect ($F(3/72) = 3.59$, $p = .018$, partial $\eta^2 = .13$) or the interaction ($F(3/72) = .16$, $p = .92$), but actually lead to a marginally significant main effect of treatment in the *opposite* direction to Experiments 1 and 2 ($F(1/24) = 3.39$, $p = .078$, partial $\eta^2 = .12$, Figure 5A). Specifically, participants showed a tendency to bid higher on shapes that were paired with incongruent than congruent trials. However, non-parametric testing of the devaluation difference scores including the outlier revealed no significant difference in bidding level between incongruent and congruent shapes (Wilcoxon's signed-rank test, $z = 1.2$; $p = .23$).

Debriefing questionnaire—None of the participants were classified as explicit or partially explicit learners of the treatment-contingency. 3 out of the 26 participants were

classified as partially explicit learners of the reward-contingencies in the learning phase. Of these, none could name the relative values of more than 4 of the shapes.

Discussion

We ran the same procedure as Experiments 1 and 2, except in the treatment phase we selectively paired one half of the shapes with incongruent Simon trials instead of stopping. Incongruent trials are more effortful (and perhaps aversive), and induced more errors and response-conflict than congruent trials. In line with our prediction, participants did not bid less for shapes that were paired with incongruent trails compared to congruent ones. Indeed, there was a trend effect in the opposite direction; participants bid slightly more for shapes paired with incongruent trials. Based on our sample size of N = 27, and an effect size of partial η^2 = .245 (which was the average effect size of the main effect of STOPPING in Experiments 1 and 2), we achieved a nominal power of greater than 99% (assuming a correlation between repeated measures of .5). Hence, even increasing power beyond our sample size would likely not reveal significant effect stimulus devaluation due to effort, error rates, or response-conflict. We interpret this experiment as suggesting that the devaluation effects of Experiments 1 and 2 were unlikely due to the effort/conflict/errors induced by the stop signal, and instead were due to the motor stopping aspect itself. However, there were still some factors that were not controlled for by the Simon procedure, such as the attentional capture and potential aversiveness of the stop-signal tones in Experiments 1 and 2. The next experiment examines these factors.

Experiment 4

Method

Participants—24 participants (mean age = 20.5y, SEM = .35, range = 18 – 24; 18 female; 3 left-handed) at UCSD provided written informed consent and received a $10 base-payment, plus money earned in the task. These were different participants from Experiments 1–3.

Materials, task, procedure, statistics—These were identical to Experiments 1 and 2, except for the treatment-phase. The treatment phase was identical to the stop-signal task in Exp. 2, with two exceptions: 1) the tone was always played at the same time as the onset of the shape, and 2) the instruction differed in that participants were not instructed to stop when they heard a tone, but instead to execute their primary response (press the response button according to the side of the screen the stimulus appeared on), and then to press another button (the space bar) on the keyboard to signal that they heard the tone. Participants were instructed to press the button according to the side of stimulus presentation as fast as possible, and then press the space bar as fast as possible after that, in case there was a tone. The second response had to be made within the inter-trial interval. The rationale for not maintaining a delay between the shape and the tone in this experiment was that infrequent signal detection after a go signal will itself induce some inhibitory control (partial recruitment of the brain's stop circuit) even if it is not a stop signal (Aron, Robbins, & Poldrack, 2014; Bissett & Logan, 2014). We reasoned that a signal at zero delay would still be equally aversive as the stop-signals in Experiments 1 and 2, and also induce attentional

capture, while not engaging motor stopping. In keeping with our rationale from Experiment 1–3, tones were displayed on half of the shapes, 75% of the time (27 of 36 trials per shape). Trial timings, trial numbers, and block numbers were identical to Experiments 1–3.

Results

Treatment phase—Mean primary-response RT for single-response trials was 358ms (SEM: 10.1), mean primary-response RT for double-response trials was 359ms (SEM: 13.1). Mean double-response success rate was 97.1% (SEM: .53; range: 90.74 – 100).

Valuation phase—The 2×4 rmANOVA revealed a significant main effect of VALUE ($F(3/69) = 5.34$, $p = .002$, partial $\eta^2 = .19$) and no main effect of TREATMENT ($F(1/23) = .13$, $p = .72$, partial $\eta^2 = .006$, Figure 5B), with no interaction ($F(3/69) = .77$, $p = .51$). No outliers with respect to the devaluation measurement were present in the sample.

Debriefing questionnaire—Two participants reported partial awareness of the treatment-contingency, but neither could name any shape that was paired more often with a tone compared to the other shapes, or alternatively, a shape that was never paired with stopping. Hence, those data remained in the sample. 7 out of the 24 participants were classified as partially explicit learners of the reward-contingencies in the learning phase. None could name the relative values of more than 4 shapes.

Discussion

We ran the same procedure as Experiments 1 and 2, except now the tones in the treatment phase did not indicate a stop-trial, but instead, a trial in which an additional response had to be made, but no response had to be stopped. Identically to the logic of Experiments 1–3, half of the shapes were paired with these double response trials and half were never paired with double-responses / tones. Double-response trials should have included the same amount of aversiveness and attentional capture as stop-signal trials (because of the presence of the tone), and should also have required more effort than single-response trials. However, they did not involve inhibitory control (primary task RT was almost identical for single- and double-response trials). In line with our predictions, participants did not bid less for shapes that were paired with the double-response requirement compared to the trials with the single-response requirement. This is in contrast to the results of Experiments 1 and 2, where stop signals did induce devaluation. Unlike Experiment 3, the numerical trend was not reversed, as participants' bids were indeed numerically lower for double-response trials. This null result is probably not due to the lack of statistical power. Assuming that double- vs. single-responses would reduce value with at least a medium effect size of partial $\eta^2 = .06$, and assuming a correlation among repeated measures of .5, then, with our sample size of N=24, we would have had a nominal achieved power of 82.5%. Actually, the effects of stopping-induced devaluation in Experiments 1 and 2 were much bigger (partial $\eta^2 = .24$ and .25, respectively). Assuming those effect sizes, our nominal power in this study exceeded 99%. Again, we interpret this experiment as suggesting that the devaluation effects of Experiments 1 and 2 were unlikely due to the aversiveness or salience of the auditory stop-signal or the increased effort associated with stopping, and instead were due to the motor stopping aspect itself.

GENERAL DISCUSSION

We developed a new behavioral paradigm to examine the effect of stopping action on value. In an initial learning phase, we associated colored geometric shapes with monetary value. In a subsequent treatment phase, we then paired half of those shapes with occasional stopping. In a final auction phase, we asked participants to bid money on the shapes in an economically valid procedure. Across two experiments, we showed that pairing stop-signals with a valuable shape decreases the subjective value of that shape. We then tested whether this devaluation effect was due to action-stopping, or whether other properties of stop-trials could account for it. Accordingly, we exchanged the stop signal task in the treatment phase with a Simon task (Experiment 3) and a double-response task (Experiment 4). These control experiments suggest that increased effort, error commission, attentional capture, aversiveness, or response-conflict cannot account for stimulus-devaluation in the absence of stopping. Instead, we argue it was specifically the action-stopping requirement of Experiments 1 and 2 that lead to stimulus devaluation.

The mechanism by which stop signals induce stimulus devaluation

Experiments 1 and 2 showed that stop signals induced stimulus devaluation, while Experiments 3 and 4 showed that incongruent Simon trials and double-response trials did not. Incongruent Simon trials are effortful and induce errors and response-conflict. Double-response trials include the same tone as the stop-signal trial, furthermore with the same frequency of occurrence, and they are more effortful than single-response trials. However, those manipulations did not induce stimulus-devaluation. Taken together, these results suggest that it was indeed the motor stopping requirement itself that induced stimulus devaluation. We further suppose that it was the successful stop trials that induced the stimulus devaluation rather than the unsuccessful ones. The main support for this view is that incongruent trials in Experiment 3, which had a higher error rate than congruent trials, did not induce stimulus devaluation (indeed, there was an increase in bidding amounts for incongruent trials). We accept that this argument is not definitive because there were not as many erroneous incongruent trials as failed stop trials, and because commission errors in a stop-signal task might be qualitatively different from direction errors in the Simon task. Ultimately, we suspect that deciding the issue of the impact of successful vs. failed stop trials on stimulus value will require neuroscience studies with the current behavioral paradigm. These will allow a trial-by-trial interrogation of value representations in the treatment phase in a way that cannot easily be done with purely behavioral methods alone. For example, behavioral methods would require a willingness-to-pay procedure on a trial-by-trial basis in the treatment phase. This is not feasible as it would make the purpose of the study overly evident to the participant, and, moreover, interspersing such evaluations would break up the prepotency of the go/stop procedure.

Successful stopping could decrease stimulus value in several ways. First, rapid successful motor stopping could induce global suppression that also impacts value. It has already been shown that successful action stopping has global suppressive effects at the *motor* level (Badry et al., 2009; Cai, Oldenkamp, & Aron, 2012; Majid, Cai, George, Verbruggen, & Aron, 2012; Wessel, Reynoso, & Aron, 2013). Accordingly, successful action stopping

could also have a suppressive effect on cognition. If value in the treatment phase is represented like working memory (Wallis, 2007) then rapid action stopping could perhaps decrement value by impacting the working memory representation. It is interesting in this regard that participants with more explicit knowledge of value in Experiments 1 and 2 showed an even larger stopping-induced devaluation effect size. A second way by which successful stopping could lead to stimulus devaluation is that repeated rapid successful motor stopping to a stimulus can attach a 'inhibitory tag' to that stimulus (Chiu, Aron, & Verbruggen, 2012; Lenartowicz et al., 2011; Verbruggen & Logan, 2008). Such an inhibitory tag can lead to slower response times upon later presentations of a 'tagged' stimulus. However, in the current paradigm participants bid less for shapes in the auction phase that had earlier been paired with stop signals, and this bidding decision was not under speed pressure, so it is not clear how an 'inhibitory tag' would lead to lower bidding. Third, rapid action-stopping could reduce value through motor/limbic 'spill-over' (Berkman et al., 2009). We expect that conclusive evidence for the mechanism underlying the stimulus devaluation effect will have to come from a neuroscience approach such as fMRI, which can measure value-, error-, orienting-, and stopping-signals in the brain. The current paradigm lends itself very well to that approach; in particular for repeatedly assaying brain representations of value during the treatment phase.

The impact of stop signals on explicit vs. implicit value representations

In our experiments, value representations were largely implicit, as confirmed by the verbal debriefing reports by the participants. Indeed, our design was aimed at making the value representations as implicit as possible, based on the intuition that the implicit valuations would be more susceptible to stopping-induced devaluation (Marteau et al., 2012). Accordingly, we designed the task to have many (eight) different shapes with four different underlying values, which increased in non-linear fashion ($.5, $1, $2, $4). Further, we included substantial dispersion around the mean reward values, and we included 20% of 'decoy' trials in which participants received no reward at all. We also ensured the learning phase was relatively short (giving less opportunity for participants to explicitly represent contingencies). However, contrary to our intuition, those participants in Experiments 1 and 2 who reported partially explicit recollections of the reward regularities showed a larger devaluation effect compared to the general sample (in terms of effect size). This finding, while largely a post-hoc observation and not the result of a systematic investigation, is interesting in several respects. First, as noted above, a greater impact of stopping on value when it is explicit fits with the possibility that stopping has its effects by decrementing memory for value representations, be these in working memory or long term memory during the treatment phase. Second, it suggests that stopping as a behavioral modification may work best with more explicit value representations. This could be a practically useful insight as the value representations of real-world objects and goods are probably at least partially explicit. Future investigations could target the question of whether explicit and implicit value representations are differentially prone to stopping-induced devaluation in a more systematic fashion.

Conclusion

We show that pairing a reward-associated stimulus with a stop-signal decreases stimulus-valuation. This provides an experimental framework within which stopping-related stimulus-devaluation can be reliably evoked, and potentially explains why action-stopping decreases the consumption of primary reinforcers, and how it can be used to effect behavioral change. Importantly, our framework consists of straightforward and easily implementable task-components, which have clear-cut operationalizations (value training, stop-signal task, auction procedure), and do not rely on having to induce thirst or hunger, or on preexisting valuations. The results from the two control experiments strongly suggest that the value reduction was due to the action-stopping rather than confounding factors of the stop signal treatment such as task demand characteristics, conflict and infrequent signal detection. While our results also suggest that the value reduction was due to successful rather than failed stopping, future studies are required to test this more thoroughly. This could best be accomplished with neuroscience approaches, which can interrogate value representations in the brain on a trial-by-trial level (Glascher, Hampton, & O'Doherty, 2009; Shao et al., 2013; Wunderlich, Rangel, & O'Doherty, 2009). In summary, this study clearly shows that pairing stop-signals with valuable stimuli leads to stimulus devaluation. This makes a novel theoretical link between the fields of action-stopping and neuroeconomics, and hints at the potential for using action-stopping as a method of behavioral intervention in clinical disorders associated with pathological stimulus overvaluation.

Acknowledgments

We thank Amanda Goold, Aiyana Bailin, and Melissa Aguilar for help with data collection, NIH NIDA R03–035874 for funding to JRW, NIH NIDA R01–026452 for funding to ARA, and two anonymous reviewers for helpful comments on earlier versions of this manuscript.

REFERENCES

Aron AR. The neural basis of inhibition in cognitive control. Neuroscientist. 2007; 13(3):214–228. [PubMed: 17519365]

Aron AR, Robbins TW, Poldrack RA. Inhibition and the right inferior frontal cortex: one decade on. Trends Cogn Sci. 2014; 18(4):177–185. [PubMed: 24440116]

Badry R, Mima T, Aso T, Nakatsuka M, Abe M, Fathi D, Fukuyama H. Suppression of human cortico-motoneuronal excitability during the Stop-signal task. Clin Neurophysiol. 2009; 120(9):1717–1723. [PubMed: 19683959]

Becker GM, Degroot MH, Marschak J. Measuring Utility by a Single-Response Sequential Method. Behavioral Science. 1964; 9(3):226–232. [PubMed: 5888778]

Berkman ET, Burklund L, Lieberman MD. Inhibitory spillover: Intentional motor inhibition produces incidental limbic inhibition via right inferior frontal cortex. Neuroimage. 2009; 47(2):705–712. [PubMed: 19426813]

Bissett PG, Logan GD. Selective stopping? Maybe not. J Exp Psychol Gen. 2014; 143(1):455–472. [PubMed: 23477668]

Botvinick M, Nystrom LE, Fissell K, Carter CS, Cohen JD. Conflict monitoring versus selection-for-action in anterior cingulate cortex. Nature. 1999; 402(6758):179–181. [PubMed: 10647008]

Brainard DH. The Psychophysics Toolbox. Spat Vis. 1997; 10(4):433–436. [PubMed: 9176952]

Cai W, Oldenkamp CL, Aron AR. Stopping speech suppresses the task-irrelevant hand. Brain Lang. 2012; 120(3):412–415. [PubMed: 22206872]

Chiu YC, Aron AR, Verbruggen F. Response suppression by automatic retrieval of stimulus-stop association: evidence from transcranial magnetic stimulation. J Cogn Neurosci. 2012; 24(9):1908–1918. [PubMed: 22624606]

Clark L, Averbeck B, Payer D, Sescousse G, Winstanley CA, Xue G. Pathological choice: the neuroscience of gambling and gambling addiction. J Neurosci. 2013; 33(45):17617–17623. [PubMed: 24198353]

Conklin CA, Tiffany S. Cue-exposure treatment: Time for change. Addiction. 2002; 97(9):1219–1221. [PubMed: 12199838]

Doallo S, Raymond JE, Shapiro KL, Kiss M, Eimer M, Nobre AC. Response inhibition results in the emotional devaluation of faces: neural correlates as revealed by fMRI. Soc Cogn Affect Neurosci. 2012; 7(6):649–659. [PubMed: 21642353]

Dutra L, Stathopoulou G, Basden SL, Leyro TM, Powers MB, Otto MW. A meta-analytic review of psychosocial interventions for substance use disorders. American Journal of Psychiatry. 2008; 165(2):179–187. [PubMed: 18198270]

Fenske MJ, Raymond JE, Kessler K, Westoby N, Tipper SP. Attentional inhibition has social-emotional consequences for unfamiliar faces. Psychological Science. 2005; 16(10):753–758. [PubMed: 16181435]

Ferrey AE, Frischen A, Fenske MJ. Hot or not: response inhibition reduces the hedonic value and motivational incentive of sexual stimuli. Front Psychol. 2012; 3:575. [PubMed: 23272002]

Fishbach A, Shah JY. Self-control in action: implicit dispositions toward goals and away from temptations. J Pers Soc Psychol. 2006; 90(5):820–832. [PubMed: 16737375]

Glascher J, Hampton AN, O'Doherty JP. Determining a role for ventromedial prefrontal cortex in encoding action-based value signals during reward-related decision making. Cereb Cortex. 2009; 19(2):483–495. [PubMed: 18550593]

Goldstein RZ, Volkow ND. Drug addiction and its underlying neurobiological basis: neuroimaging evidence for the involvement of the frontal cortex. Am J Psychiatry. 2002; 159(10):1642–1652. [PubMed: 12359667]

Goldstein RZ, Volkow ND. Dysfunction of the prefrontal cortex in addiction: neuroimaging findings and clinical implications. Nature Reviews Neuroscience. 2011; 12(11):652–669.

Hammersley R. Cue Exposure and Learning-Theory. Addictive Behaviors. 1992; 17(3):297–300. [PubMed: 1353283]

Heather N, Bradley BP. Cue Exposure as a Practical Treatment for Addictive Disorders - Why Are We Waiting. Addictive Behaviors. 1990; 15(4):335–337. [PubMed: 2248107]

Houben K. Overcoming the urge to splurge: Influencing eating behavior by manipulating inhibitory control. Journal of Behavior Therapy and Experimental Psychiatry. 2011; 42(3):384–388. [PubMed: 21450264]

Houben K, Havermans RC, Nederkoorn C, Jansen A. Beer a no-go: learning to stop responding to alcohol cues reduces alcohol intake via reduced affective associations rather than increased response inhibition. Addiction. 2012; 107(7):1280–1287. [PubMed: 22296168]

Houben K, Jansen A. Training inhibitory control. A recipe for resisting sweet temptations. Appetite. 2011; 56(2):345–349. [PubMed: 21185896]

Houben K, Wiers RW, Jansen A. Getting a Grip on Drinking Behavior: Training Working Memory to Reduce Alcohol Abuse. Psychological Science. 2011; 22(7):968–975. [PubMed: 21685380]

Izuma K, Matsumoto M, Murayama K, Samejima K, Sadato N, Matsumoto K. Neural correlates of cognitive dissonance and choice-induced preference change. Proc Natl Acad Sci U S A. 2010; 107(51):22014–22019. [PubMed: 21135218]

Kiss M, Goolsby BA, Raymond JE, Shapiro KL, Silvert L, Nobre AC, Eimer M. Efficient attentional selection predicts distractor devaluation: event-related potential evidence for a direct link between attention and emotion. J Cogn Neurosci. 2007; 19(8):1316–1322. [PubMed: 17651005]

Lenartowicz A, Verbruggen F, Logan GD, Poldrack RA. Inhibition-related activation in the right inferior frontal gyrus in the absence of inhibitory cues. J Cogn Neurosci. 2011; 23(11):3388–3399. [PubMed: 21452946]

Logan GD, Cowan WB, Davis KA. On the Ability to Inhibit Simple and Choice Reaction-Time Responses - a Model and a Method. Journal of Experimental Psychology-Human Perception and Performance. 1984; 10(2):276–291. [PubMed: 6232345]

Majid DS, Cai W, George JS, Verbruggen F, Aron AR. Transcranial magnetic stimulation reveals dissociable mechanisms for global versus selective corticomotor suppression underlying the stopping of action. Cereb Cortex. 2012; 22(2):363–371. [PubMed: 21666129]

Marteau TM, Hollands GJ, Fletcher PC. Changing human behavior to prevent disease: the importance of targeting automatic processes. Science. 2012; 337(6101):1492–1495. [PubMed: 22997327]

Phaf RH, Rotteveel M. Affective monitoring: a generic mechanism for affect elicitation. Front Psychol. 2012; 3:47. [PubMed: 22403557]

Rothemund Y, Preuschhof C, Bohner G, Bauknecht HC, Klingebiel R, Flor H, Klapp BF. Differential activation of the dorsal striatum by high-calorie visual food stimuli in obese individuals. Neuroimage. 2007; 37(2):410–421. [PubMed: 17566768]

Shao R, Read J, Behrens TE, Rogers RD. Shifts in reinforcement signalling while playing slot-machines as a function of prior experience and impulsivity. Transl Psychiatry. 2013; 3:e235. [PubMed: 23443361]

Simon JR, Rudell AP. Auditory S-R Compatibility - Effect of an Irrelevant Cue on Information Processing. Journal of Applied Psychology. 1967; 51(3):300. [PubMed: 6045637]

Stevens, James. Applied multivariate statistics for the social sciences. Taylor & Francis US; 2009.

Veling H, Aarts H, Papies EK. Using stop signals to inhibit chronic dieters' responses toward palatable foods. Behaviour Research and Therapy. 2011; 49(11):771–780. [PubMed: 21906724]

Veling H, Aarts H, Stroebe W. Stop signals decrease choices for palatable foods through decreased food evaluation. Front Psychol. 2013a; 4:875. [PubMed: 24324451]

Veling H, Aarts H, Stroebe W. Using stop signals to reduce impulsive choices for palatable unhealthy foods. British Journal of Health Psychology. 2013b; 18(2):354–368. [PubMed: 23017096]

Verbruggen F, Adams R, Chambers CD. Proactive Motor Control Reduces Monetary Risk Taking in Gambling. Psychological Science. 2012; 23(7):805–815. [PubMed: 22692336]

Verbruggen F, Aron AR, Stevens MA, Chambers CD. Theta burst stimulation dissociates attention and action updating in human inferior frontal cortex. Proc Natl Acad Sci U S A. 2010

Verbruggen F, Logan GD. Automatic and Controlled Response Inhibition: Associative Learning in the Go/No-Go and Stop-Signal Paradigms. Journal of Experimental Psychology-General. 2008; 137(4):649–672. [PubMed: 18999358]

Verbruggen F, Logan GD. Models of response inhibition in the stop-signal and stop-change paradigms. Neurosci Biobehav Rev. 2009; 33(5):647–661. [PubMed: 18822313]

Wallis JD. Orbitofrontal cortex and its contribution to decision-making. Annual Review of Neuroscience. 2007; 30:31–56.

Wessel, JR.; Huber, DE.; Aron, AR. Rapidly stopping action leads to forgetting; Paper presented at the Society for Neuroscience; New Orleans. 2012.

Wessel JR, Reynoso HS, Aron AR. Saccade suppression exerts global effects on the motor system. J Neurophysiol. 2013

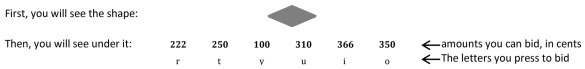
Wiers RW, Eberl C, Rinck M, Becker ES, Lindenmeyer J. Retraining automatic action tendencies changes alcoholic patients' approach bias for alcohol and improves treatment outcome. Psychological Science. 2011; 22(4):490–497. [PubMed: 21389338]

Wunderlich K, Rangel A, O'Doherty JP. Neural computations underlying action-based decision making in the human brain. Proc Natl Acad Sci U S A. 2009; 106(40):17199–17204. [PubMed: 19805082]

APPENDIX 1: AUCTION INSTRUCTIONS

You have already earned a show-up fee of $10 dollars. You've also made some money based on your performance in the first task. Now, you have the opportunity to make even more money in this auction phase.

We will show you different polygon symbols, the same ones you've been seeing all along. Each time you see a symbol, you will place a bid for that symbol.



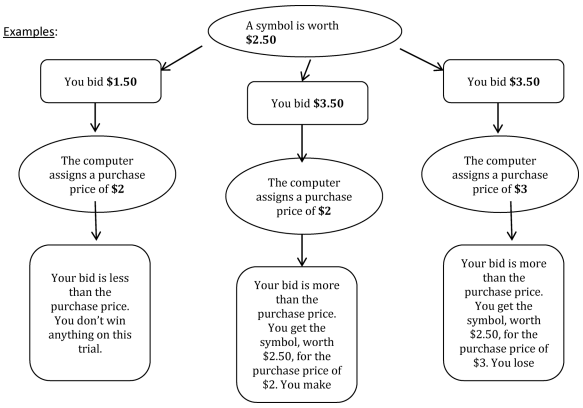
You will perform multiple trials, each time making a bid for the symbol, but only five randomly selected trials will count. Since you don't know which trials count, you should treat every decision as though it is the only one.

Definitions

- **True Value**: The monetary amount that a symbol is actually worth. If you win the symbol, you will be PAID this amount in real money, minus what you paid for the symbol (the purchase price). If you bid wisely, you will be able to buy symbols for less than their true value, and make a profit.
- **Purchase Price**: An amount of money, between $0 and $4, chosen randomly by the computer after you place your bid. If your bid is greater than or equal to the purchase price, you will get to "buy" that symbol at the purchase price. If your bid is less than the purchase price, you do not get anything.

Note

- The symbols are associated with different amounts (their true values).
- You will have $4 dollars at your disposal on every new trial – so you can bid $4 max.



Strategy

The best thing you can do is to always bid the number that is CLOSEST to your true valuation for a given symbol. Every trial you should ask yourself how much of the $4 dollars you want to spend on that particular symbol.

Common mistakes

You might think that your best strategy is to *bid less than the symbol is worth to you*. This is INCORRECT. Note that the price that you pay is determined randomly by the computer (the purchase price) and NOT by your bid. Thus, by lowering your bid you will not be able to affect the price that you pay, but you might end up losing the opportunity to buy the symbol at a "good" price. This is show in Example 1 above. There, the bet, at $1.50 was too low.

You also might think that, since you have $4 to spend every time, you should *always bid the highest amount* to win the most symbols. This is also INCORRECT. If you do this, you will wind up paying more for some symbols than they are worth, and lose money. This is shown in Example 3 above. There, the bet at $3.50 was too high.

Also, some people chose to utilize a strategy in which they *bet very similar amounts, regardless of which symbol they are bidding fo*r (say, e.g. they pick the value closest to $2 / 200c on each trial) in order to minimize their possible error margin across symbols. This is actually the WORST possible technique, since this will lead to under-bidding on valuable symbols (resulting in not winning the auction on these symbols) and over-bidding on the least valuable symbols (resulting in 'winning' the auction for that symbol, but at way too high prizes).

Quiz

Imagine the true value of a symbol is $2. Which amount should you bid?

A. $1

B. $2

C. $3

In which one of the following conditions will you get the symbol?

A. When you bid 2$ and a purchase price of $1 is drawn

B. When you bid 2$ and a purchase price of $3 is drawn

Reminder/clarification

- You will make a bid for "purchasing" a symbol that could be worth more than the values you can bet (>$4).
- You will bid for purchasing symbols on multiple trials, but only five randomly selected trials will count.
- You should treat each decision as if it were the only one, and independent from all the others.
- Your best strategy is to always bid the number closest to your true value for that symbol.

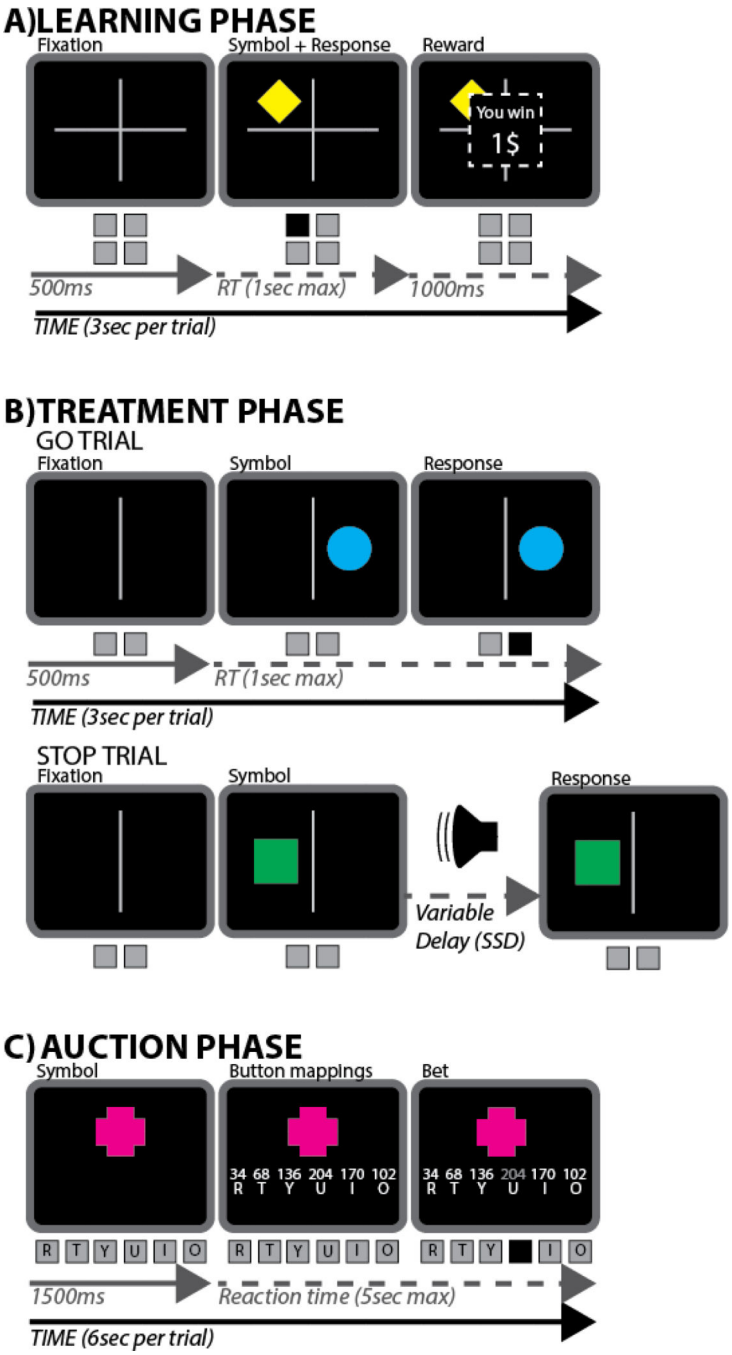


Figure 1.
Task diagram for Experiments 1 and 2. A) Learning phase. B) Treatment phase. C) Auction phase.

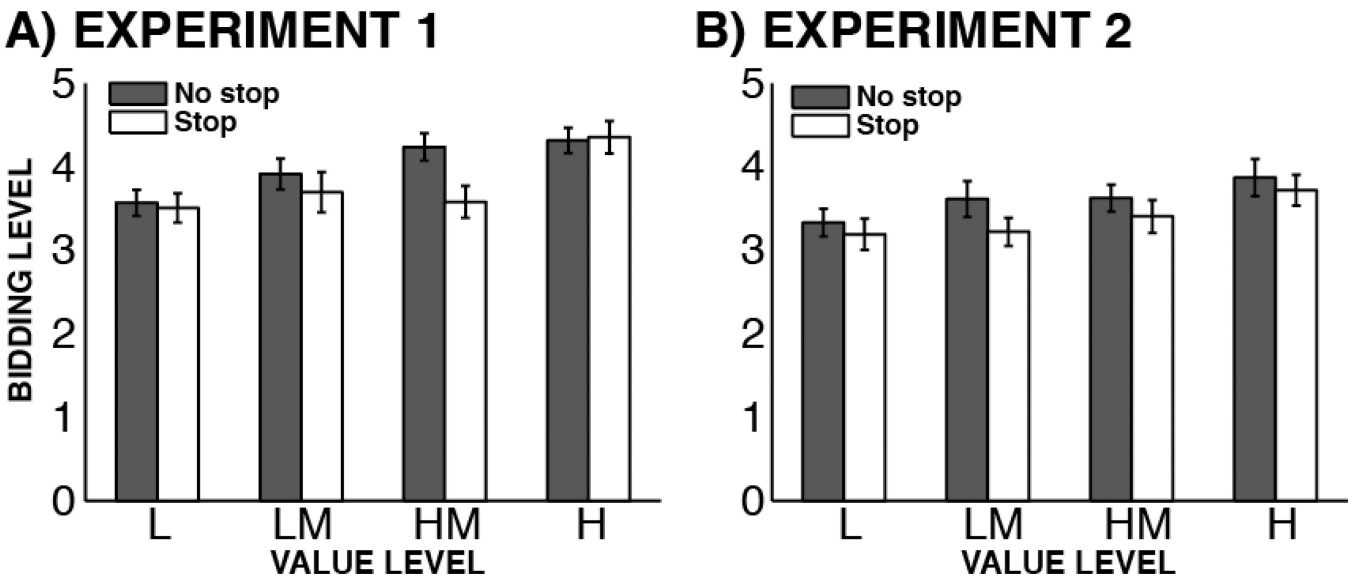


Figure 2.

Bidding data from the valuation phase in Experiments 1 and 2. Depicted are bidding levels from low bids (1) to high bids (6), plotted by the actual value of each shape from the learning phase (L = low, LM = low-medium, HM = high-medium, H = high). No stop and Stop refer to whether the shapes were paired stop-signals or not in the earlier treatment phase. Error bars denote the standard error of the mean across subjects.

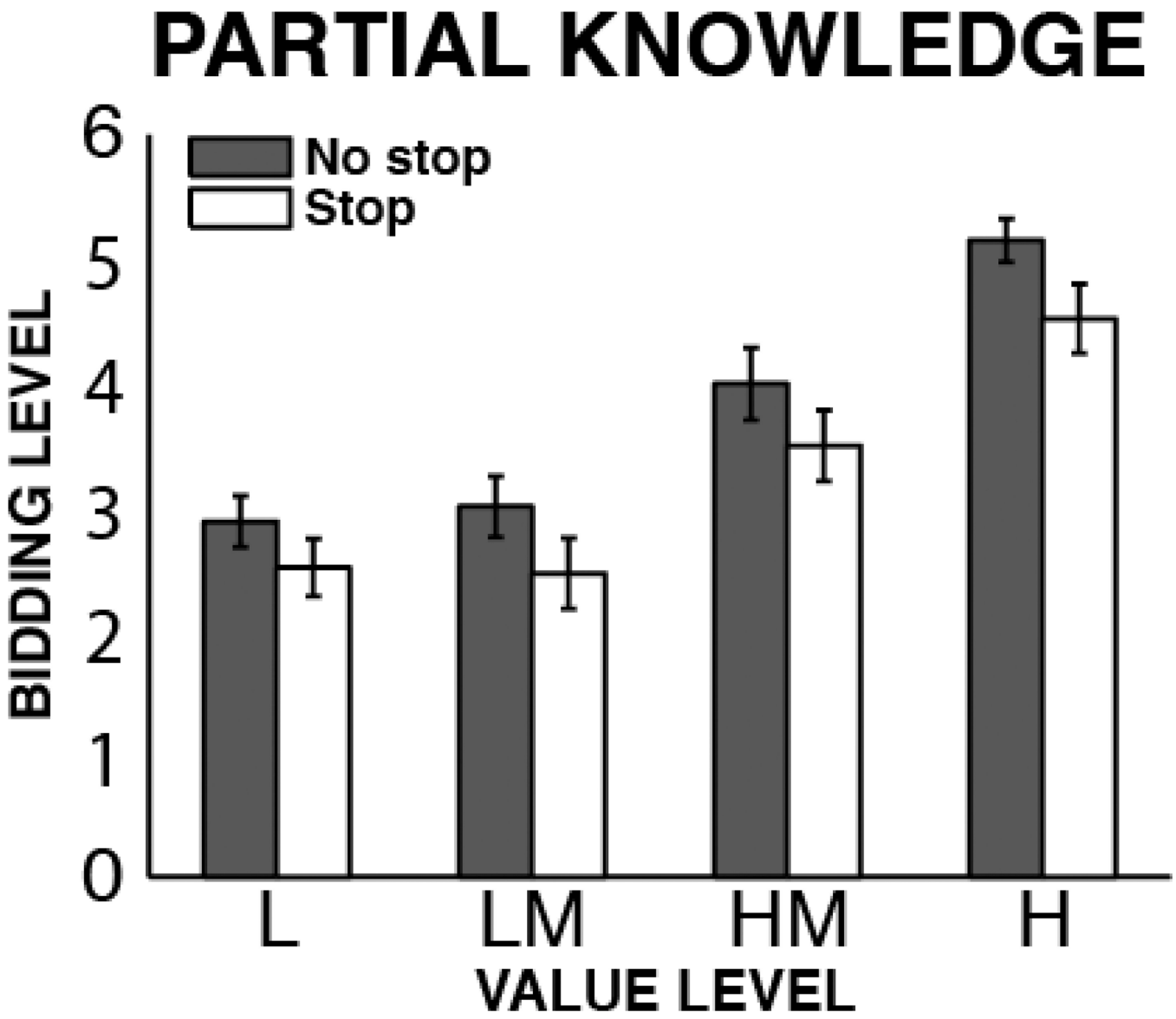


Figure 3.

Bidding data from the valuation phase in Experiments 1 and 2; pooled data from N = 15 participants that reported partial awareness of the reward contingencies of the task. Depicted are bidding levels from low bids (1) to high bids (6), plotted by the actual value of each shape from the learning phase (L = low, LM = low-medium, HM = high-medium, H = high). Error bars denote the standard error of the mean across subjects.

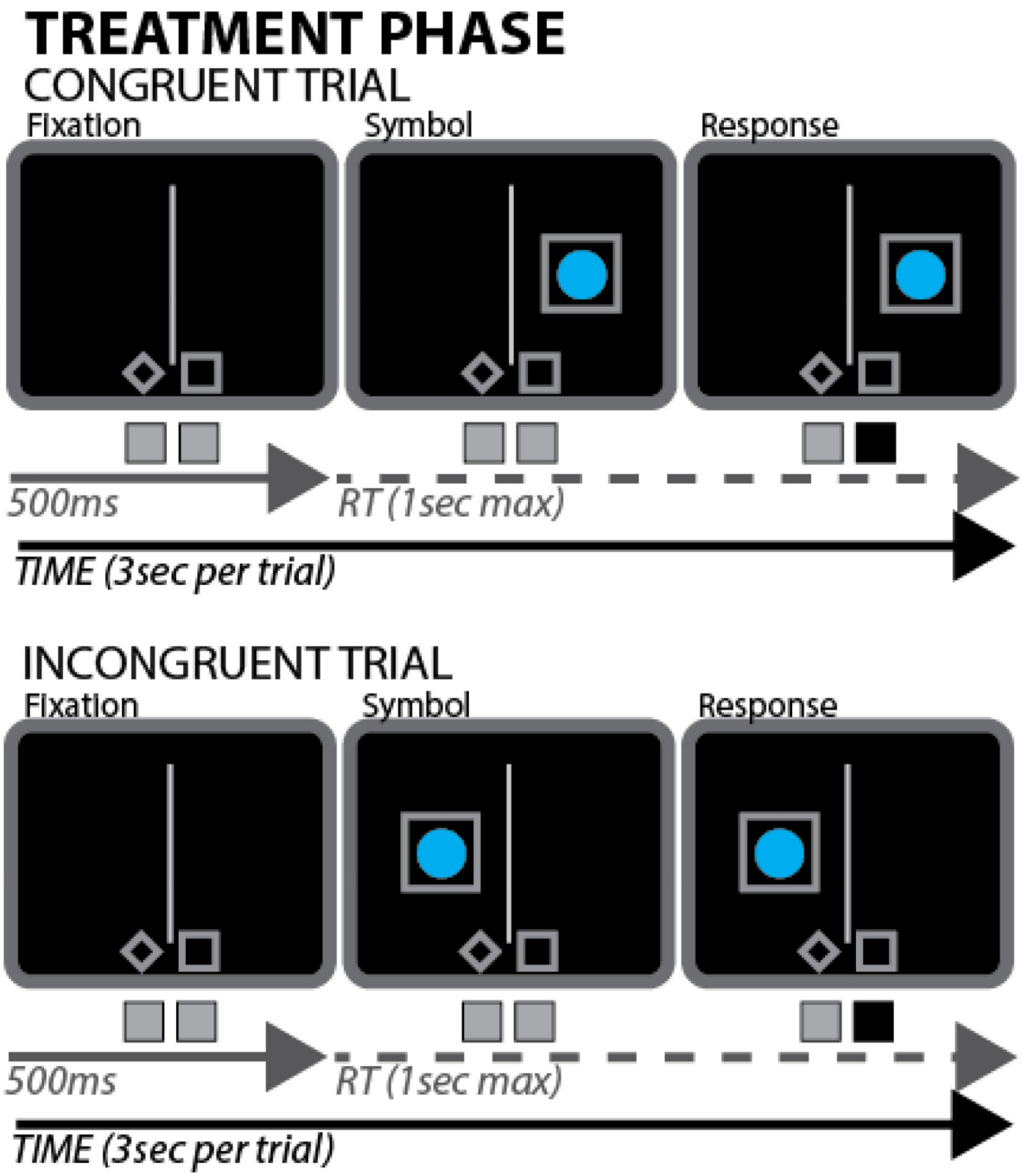


Figure 4.
Trial diagram for the treatment phase in Experiment 3. On each trial, participants had to respond quickly according to the outer polygon, i.e., the square or diamond that enclosed the valuable shape.

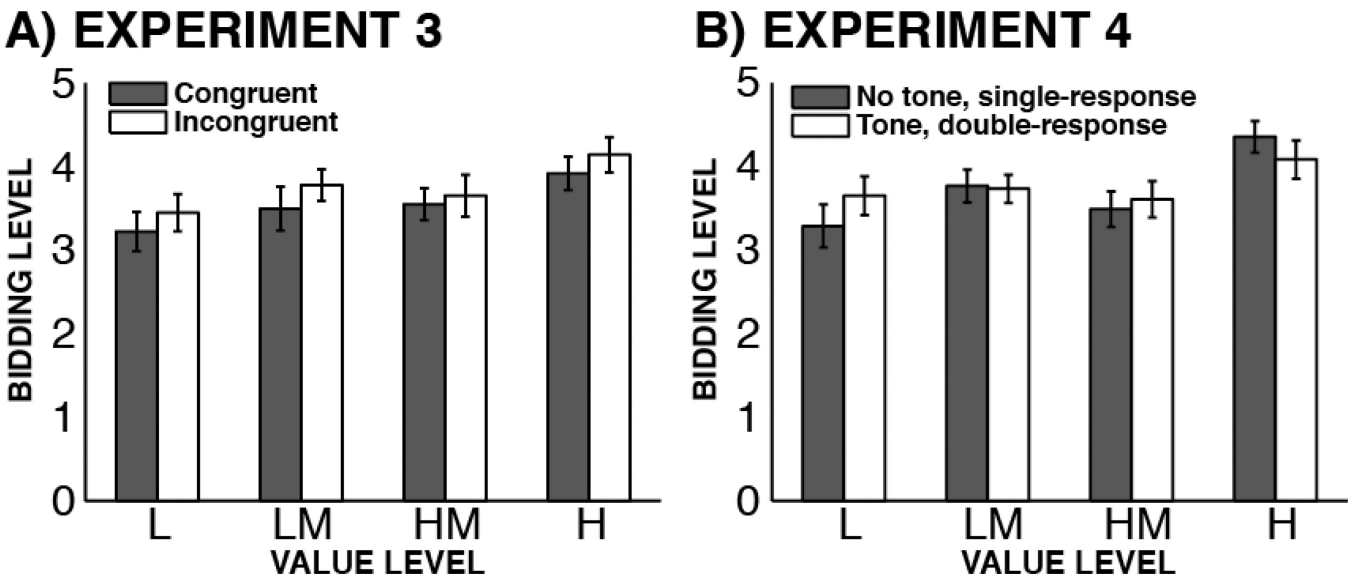


Figure 5.

Bidding data from the valuation phase in Experiments 3 and 4. Depicted are bidding levels from low bids (1) to high bids (6), plotted by the actual value of each shape from the learning phase (L = low, LM = low-medium, HM = high-medium, H = high). A) Congruent denotes trials that were paired with congruent Simon trials only (in the earlier treatment phase), incongruent denotes trials that were paired with incongruent Simon stimuli on 75% of trials. B) No-tone, single-response denotes trials that were always paired with the single-response requirement, Tone, double-response denotes trials that were paired with the double-response requirement 75% of the time. Error bars denote the standard error of the mean across subjects.